D Cepstrum

D.1 Echoes

The famous paper with the silly words is by Bogert et al. (1963). It's quite a heuristic study motivated by seismic echoes, but their theoretical motivation is as follows: Say there is a signal $x(t)$ added to a delayed and scaled version of itself

$$y(t) = x(t) + \alpha x(t - \tau). \tag{134}$$

The Fourier transform and power spectrum of that signal are then

$$Y(f) = X(f) + \alpha X(f) e^{j2\pi f \tau} \tag{135}$$

$$Y(f)^2 = X(f)^2 (1 + 2\alpha \cos 2\pi f \tau + \alpha^2). \tag{136}$$

Given that $\alpha < 1$, and

$$\log(1 + x) = \sum_{n=1}^{\infty} \frac{x^n}{n} \tag{137}$$

the log power spectrum can be approximated as

$$2 \log Y(f) = 2 \log X(f) + 2\alpha \cos 2\pi f \tau. \tag{138}$$

So, the echo manifests itself as ripple in the log spectrum. It then makes sense to take the Fourier transform of this to analyse the ripple.

D.2 Oppenheim's approach

The approach of Oppenheim and Schafer (1968, 1989) is to say that, under DFT or z-transform, convolutionally mixed signals are multiplicative. Under $\log F(z)$ they would be additive. $F(z)$ is complex, so that's a complex logarithm.

That's already homomorphic; it doesn't yet justify another transform. The log of a sum isn't easily tractable, but assume that the z-transform has the form of a ratio of polynomials,

$$F(z) = \frac{\prod_{k=1}^{M} (1 - a_k z^{-1})}{\prod_{k=1}^{N} (1 - c_k z^{-1})} \tag{139}$$

$$= \frac{\prod_{k=1}^{M_i} (1 - a_k z^{-1}) \prod_{k=1}^{M_o} (1 - b_k z)}{\prod_{k=1}^{N_i} (1 - c_k z^{-1}) \prod_{k=1}^{N_o} (1 - d_k z)} A z^r \tag{140}$$

where the second line distinguishes poles and zeros inside the unit circle from those outside, i.e., if $a_k > 1$ then write $b_k = 1/a_k$. Assume the time series can be scaled and shifted such that $A > 0$ and z^r can be ignored. Given

$$\log(1 - \alpha z^{-1}) = -\sum_{n=1}^{\infty} \frac{\alpha^n}{n} z^{-n}, \tag{141}$$

$$\log(1 - \beta z) = -\sum_{n=1}^{\infty} \frac{\beta^n}{n} z^{n}, \tag{142}$$

we can then write

$$\log F(z) = \sum_{k=1}^{M_i} \log(1 - a_k z^{-1}) + \sum_{k=1}^{M_o} \log(1 - b_k z) - \sum_{k=1}^{N_i} \log(1 - c_k z^{-1}) - \sum_{k=1}^{N_o} \log(1 - d_k z) + \log(A) \tag{143}$$

$$= -\sum_{k=1}^{M_i} \sum_{n=1}^{\infty} \frac{a_k^n}{n} z^{-n} - \sum_{k=1}^{M_o} \sum_{n=1}^{\infty} \frac{b_k^n}{n} z^{n} + \sum_{k=1}^{N_i} \sum_{n=1}^{\infty} \frac{c_k^n}{n} z^{-n} + \sum_{k=1}^{N_o} \sum_{n=1}^{\infty} \frac{d_k^n}{n} z^{n} + \log(A) \tag{144}$$

$$= \sum_{n=-\infty}^{-1} \left[\sum_{k=1}^{M_o} \frac{b_k^{-n}}{n} z^{-n} - \sum_{k=1}^{N_o} \frac{d_k^{-n}}{n} z^{-n} \right] + \log(A) + \sum_{n=1}^{\infty} \left[-\sum_{k=1}^{M_i} \frac{a_k^n}{n} z^{-n} + \sum_{k=1}^{N_i} \frac{c_k^n}{n} z^{-n} \right]. \tag{145}$$

That clearly has the form of a z-transform, the inverse of which is

$$x_n = \begin{cases} \sum_{k=1}^{M_o} \frac{b_k^{-n}}{n} - \sum_{k=1}^{N_o} \frac{d_k^{-n}}{n} & n < 0, \\ \log(A) & n = 0, \\ -\sum_{k=1}^{M_i} \frac{a_k^n}{n} + \sum_{k=1}^{N_i} \frac{c_k^n}{n} & n > 0. \end{cases} \tag{146}$$

The upshot of that is that the inverse z-transform of $\log F(z)$ separates causal and anticausal components of the signal. So, we have another definition of the (complex) cepstrum:

$$c_n = \frac{1}{2\pi j} \oint_C \log F(z) z^{n-1} \, dz. \tag{147}$$

Equation 147 defines the cepstrum of the sequence x.

D.2.1 The Fourier trick

The last trick is that that contour integral can be over any contour in the region of convergence. The one that makes sense is the unit circle; this mean we can use the discrete time Fourier transform.

$$F(\omega) = \sum_{n=-\infty}^{\infty} x_n e^{-j\omega n} \tag{148}$$

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log F(\omega) e^{j\omega n} \, d\omega. \tag{149}$$

In practice, it's necessary to use the DFT. But notice that's not the definition, it's just the means of application.

References

B. P. Bogert, M. J. R. Healy, and J. W. Tukey. The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking. In M. Rosenblatt, editor, *Proceedings of Symposium on Time Series Analysis*, pages 209–243. John Wiley and Sons, New York, 1963.

Alan V. Oppenheim and Ronald W. Schafer. Homomorphic analysis of speech. *IEEE Transactions on Audio and Electroacoustics*, AU-16(2):221–226, June 1968.

Alan V. Oppenheim and Ronald W. Schafer. *Discrete Time Signal Processing*. Signal Processing Series. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 07632, 1989.